Simultaneous zeros of a Cubic and Quadratic form

Jahan Zahid

1 Introduction

Consider a system of forms

$$\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), \ldots, F_r(\mathbf{x}))$$

of degrees $d_1, \ldots, d_r$ respectively in the variables $\mathbf{x} = (x_1, \ldots, x_n)$ over a $\mathfrak{p}$-adic field K. It had been conjectured by Artin [1, Preface] that $\mathbf{F}$ necessarily has a non-trivial zero in K provided $n > \sum d_i^2$. It should be noted that there exist systems with $n = \sum d_i^2$ which have only the trivial $\mathfrak{p}$-adic zero, so this is the best we can hope for.

Artin's conjecture has been verified in only a handful of cases. For example it is a classical result due to Hasse [8] that every quadratic form with $n > 4$ variables has a non-trivial $\mathfrak{p}$-adic zero. The case of cubic forms was settled independently by Dem'yanov [6], Lewis [10] and Springer [14]. Dem'yanov [7] and later Birch, Lewis & Murphy [3] proved the conjecture for a system of two quadratic forms. The purpose of this paper is to prove the next case of the conjecture for two forms, provided we have a large enough residue class field. More precisely we shall prove

Theorem 1. *Any system of a cubic and quadratic form in at least* 14 *variables defined over* K*, has a non-trivial zero in* K *provided the cardinality of the residue class field exceeds* 293.

It should be noted that Artin's conjecture was shown to be false in general by Terjanian [15], who found a counterexample of a quartic form in 18 variables with no zero in $\mathbb{Q}_2$. If however $|K : \mathbb{Q}_p| = e$ and $\mathbf{d} = (d_1, \ldots, d_r)$, then by a remarkable theorem of Ax & Kochen [2], there exists an integer $p(\mathbf{d}, e)$ such that any system $\mathbf{F}$ with $n > \sum d_i^2$ has a non-trivial zero provided the characteristic of the residue class field exceeds $p(\mathbf{d}, e)$.

We remark that Theorem 1 is stronger than anything we can deduce from the Ax–Kochen theorem for a number of reasons. Firstly we have an explicit bound on the cardinality of the residue class field for which Artin's conjecture is true. Secondly we have a condition depending on the cardinality of the residue class field, rather than the characteristic. Consequently we are now able to say that Artin's conjecture holds for a cubic and quadratic form over any unramified extension of $\mathbb{Q}_p$ of degree at least 9. Where it was not possible to make this deduction before.

As an outline to prove Theorem 1 we shall generalise a $\mathfrak{p}$-adic minimization procedure due to Schmidt [13] to hold for systems of forms of arbitrary degrees.

We shall then derive some Geometric information of the system over the residue class field, for those systems which terminate in the minimization process. This will allow us to find a non-singular zero in the residue class field to which we can apply Hensel's Lemma.

Acknowledgments: This work forms part of the authors doctoral thesis at the University of Oxford. I very gratefully acknowledge the financial support I received from EPSRC.

I would also like to thank my supervisor Prof. Roger Heath-Brown, for suggesting this problem and his excellent guidance over the last few years. I have also benefited from numerous conversations with Dr. Damiano Testa and Prof. Trevor Wooley, to whom I am glad to express my gratitude.

2 Some preliminaries

Let $\mathcal{O}_K$ denote the ring of integers of K, and denote the residue class field by $\mathbb{F}_q$. Let π denote a uniformizer for $\mathcal{O}_K$. If $\alpha \in K - \{0\}$, we may write $\alpha = \pi^s u$, where u is a unit in $\mathcal{O}_K$. We define the π-adic order $v(\cdot)$ by setting $v(\alpha) = s$. We also define the π-adic valuation $|\cdot|$ by setting $|\alpha| = p^{-s}$, where p denotes the characteristic of the residue class field $\mathbb{F}_q$. Recall that

$$F_1(\mathbf{x}), \ldots, F_r(\mathbf{x}) \in K[\mathbf{x}]$$

denotes an arbitrary system of forms of degrees $d_1, \ldots, d_r$ in the variables $\mathbf{x} = (x_1, \ldots, x_n)$ over K. We shall assume that $d_1 \geq \ldots \geq d_r$ unless we state otherwise and for brevity write the above system of forms as $\mathbf{F}$. We are interested in determining the existence of a point $\mathbf{x} \in K^n - \{\mathbf{0}\}$ such that $\mathbf{F}(\mathbf{x}) = \mathbf{0}$. Clearly we may assume that the coefficients of the forms $\mathbf{F}$ and the variables $\mathbf{x}$ are in $\mathcal{O}_K$, since this does not affect existence of a zero.

By a slight abuse of notation we write $GL(n, \mathcal{O}_K)$ to denote the set of $(n \times n)$-matrices over $\mathcal{O}_K$ with non-zero (rather than unitary) determinant. So let $\tau \in GL(n, \mathcal{O}_K)$ and write $\mathbf{F}_\tau$ to denote $\mathbf{F}(\tau\mathbf{x}) = (F_1(\tau\mathbf{x}), \ldots, F_r(\tau\mathbf{x}))^t$. We also write $T = (t_{ij})$ to denote the $(r \times r)$ upper triangular matrix with entries

$$t_{ij}(\mathbf{x}) = \pi^{-c_i} G_{ij}(\mathbf{x}), \tag{1}$$

where $c_i \geq 0$ and $G_{ij} \in \mathcal{O}_K[\mathbf{x}]$ to denote an arbitrary form of degree

$$\deg F_i - \deg F_j \geq 0,$$

for $1 \leq i < j \leq r$ and define the diagonal terms $G_{ii} := 1$.

Let $\Omega = (\omega_1, \ldots, \omega_r)$ where $\omega_i > 0$ for each i. Then we write

$$\mathbf{F} \underset{\Omega}{\succ} \mathbf{F}',$$

if $\mathbf{F}$ and $\mathbf{F}'$ are both defined over $\mathcal{O}_K$ and

$$\mathbf{F}' = T\mathbf{F}_\tau$$

with

$$\sum c_i \omega_i - s > 0 \tag{2}$$

where $s = v(\det \tau)$ and the c_i are as in (1). If $\mathbf{F}' = T\mathbf{F}_\tau$, it is clear that if $\mathbf{F}'$ has a zero if and only if $\mathbf{F}$ has a zero.

We say that $\mathbf{F}$ is Ω-bottomless if there is an infinite chain

$$\mathbf{F} \underset{\Omega}{\succ} \mathbf{F}^{(1)} \underset{\Omega}{\succ} \mathbf{F}^{(2)} \underset{\Omega}{\succ} \cdots$$

otherwise, $\mathbf{F}$ will be called Ω-bottomed. We also say that $\mathbf{F}$ is Ω-reduced if there does not exist any $\mathbf{F}'$ such that

$$\mathbf{F} \underset{\Omega}{\succ} \mathbf{F}'.$$

We say that two systems $\mathbf{F}$ and $\mathbf{F}'$ are equivalent, if both systems are defined over $\mathcal{O}_K$ and

$$\mathbf{F}' = T\mathbf{F}_\tau$$

where $c_i = 0$ for all $1 \le i \le r$ and $v(\det \tau) = 0$ in T and τ as in (1). The order $o(\mathbf{F})$, of a system $\mathbf{F}$ is the least positive integer m such that $\mathbf{F}$ is equivalent to a system that contains m variables explicitly. We also define the h-invariant for a system $\mathbf{F}$, denoted $h(\mathbf{F})$ as the least integer h such that we can write

$$F_i(\mathbf{x}) = x_1 H_{i1}(\mathbf{x}) + \ldots + x_h H_{ih}(\mathbf{x}) \pmod{\pi}, \tag{3}$$

for all $1 \le i \le r$ and all systems equivalent to $\mathbf{F}$. Note that since the F_i are defined over $\mathcal{O}_K$, considering them modulo π is well defined.

Given any set $S = \{e_1, \ldots, e_s\}$ of positive integers we define v_S as the least integer v such that every system consisting of s forms of degrees $e_1, \ldots, e_s$ have a non-trivial zero provided the number of variables in the system is at least v_S. If ϕ denotes the empty set we define $v_\phi := 1$. We remark here that it is due to a classical theorem of Brauer [4], the number v_S is always finite.

Although the next theorem is likely to have further applications, it will for the purpose of this paper play a crucial part in the minimisation procedure for a system of a cubic and quadratic form.

Theorem 2. *Let $S \subset \{d_1, \ldots, d_r\}$ denote any subset of cardinality $r-1$ with indexing set I such that v_S is maximal. Let $j \notin I$ then provided*

$$n \ge v_S + d_j^2 \tag{4}$$

there exists some $\Omega = (\omega_1, \ldots, \omega_r)$ such that $\omega_i > d_i$ for each $1 \le i \le r$ and such that every Ω-bottomless system $\mathbf{F}$ defined over $\mathcal{O}_K$ has a non-trivial $\mathfrak{p}$-adic zero.

3 Proof of Theorem 2

Since the field K has characteristic 0, given any form F of degree d there is a unique form $M_F(\mathbf{x}_1, \ldots, \mathbf{x}_d)$ which is linear in each vector $\mathbf{x}_j$ and which is symmetric in $\mathbf{x}_1, \ldots, \mathbf{x}_d$, such that

$$F(\mathbf{x}) = M_F(\mathbf{x}, \ldots, \mathbf{x}).$$

Let $\mathbf{e}_1 = (1, 0, \ldots, 0), \mathbf{e}_2, \ldots, \mathbf{e}_n$ be unit vectors. We say that $\mathbf{F} = (F_1, \ldots, F_r)$ is Ω-special if there are non-negative integers $a_1, \ldots, a_n$ and $b_1, \ldots, b_r$ with

$$a_1 + \cdots + a_n < \omega_1 b_1 + \cdots + \omega_r b_r \tag{5}$$

such that

$$M_{F_i}(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_{d_i}}) = 0 \tag{6}$$

for each $1 \leq i \leq r$ and d_i-tuple $(j_1, \ldots, j_{d_i})$ for which,

$$a_{j_1} + \cdots + a_{j_{d_i}} < b_i. \tag{7}$$

Note the following important correspondence between Ω-bottomless systems and Ω-special systems.

Theorem 3. *Every Ω-bottomless system is equivalent to a Ω-special system.*

We prove this in due course. Now we make use of this to prove Theorem 2.

Proof of Theorem . By Theorem 3 we may suppose that $\mathbf{F}$ is Ω-special. For ease of notation we may assume that

$$a_1 \leq \ldots \leq a_n \quad \text{and} \quad \frac{b_1}{d_1} \leq \ldots \leq \frac{b_r}{d_r}, \tag{8}$$

dropping any previous ordering we had on $d_1, \ldots, d_r$. If there is a subset $S \subset \{d_1, \ldots, d_r\}$ with indexing set I such that

$$d_i a_{v_S} < b_i \quad \text{for all } \ i \notin I$$

then $\mathbf{F}$ has a non-trivial zero. For if such a subset S exists then by (7) one has

$$M_{F_i}(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_{d_i}}) = 0 \quad \text{for all } \ i \notin I$$

and every $1 \leq j_1, \ldots, j_{d_i} \leq v_S$. Therefore the system $(F_i)_{i \notin I}$ vanishes on the v_S-dimensional subspace spanned by $\{\mathbf{e}_1, \ldots, \mathbf{e}_{v_S}\}$ and on this subspace we can find a zero of $(F_i)_{i \in I}$ and therefore a zero of $\mathbf{F}$. Consequently we may assume that for each $S \subset \{d_1, \ldots, d_r\}$ there exists some i, which by the ordering (8) we may assume to be $\min\{1 \leq i \leq d : i \notin I\}$ such that

$$d_i a_{v_S} \geq b_i.$$

We define $S_0 := \phi$ and $S_i := \{d_1, \ldots, d_i\}$ for $1 \leq i \leq r-1$ and for ease of notation write $w_i = v_{S_{i-1}}$ for $1 \leq i \leq r$. Then it follows that

$$d_i a_{w_i} \geq b_i \tag{9}$$

for every $1 \leq i \leq r$. Note that by assumption (4) we have that

$$n \geq w_r + d_r^2 \geq d_1^2 + \cdots + d_r^2 + 1.$$

Moreover for each $0 \leq i \leq r-1$ we claim that

$$w_{i+1} - w_i \geq d_i^2. \tag{10}$$

For if we let $\mathbf{T} = (F_1, \ldots, F_{i-1})$ denote a system with $w_i - 1$ variables with only the trivial zero and F_i denote a form in d_i^2 variables with only the trivial zero with its variables distinct from the variables in $\mathbf{T}$ then it is clear that the system

$\mathbf{T} \cup (F_i)$ has only the trivial zero. Therefore it follows that $w_{i+1} \geq w_i + d_i^2$ as claimed. By (8), (9) and (10) it follows that

$$\begin{aligned} a_1 + \cdots + a_n &\geq d_1^2 a_{w_1} + d_2^2 a_{w_2} + \cdots + d_r^2 a_{w_r} + a_n \\ &\geq \left(d_1^2 + \frac{1}{r}\right) a_{w_1} + \left(d_2^2 + \frac{1}{r}\right) a_{w_2} + \cdots + \left(d_r^2 + \frac{1}{r}\right) a_{w_r} \\ &\geq (d_1 + \epsilon) b_1 + (d_2 + \epsilon) b_2 + \cdots + (d_r + \epsilon) b_r \end{aligned}$$

taking $\epsilon = (r d_{max})^{-1}$, where $d_{max} := \max\{d_1, \ldots, d_r\}$. Hence if we let $\omega_i = d_i + \epsilon$ for $1 \leq i \leq r$, it follows that every Ω-special system must have a non-trivial zero completing the proof of the theorem. ☐

We shall now proceed by proving Theorem 3, generalising where appropriate the method of Schmidt [13]. Given two systems $\mathbf{F}$ and $\mathbf{F}'$ defined over $\mathcal{O}_K$, we write

$$\mathbf{F} \underset{\Omega}{\overset{k}{\succ}} \mathbf{F}' \tag{11}$$

if $k \geq 1$ and

$$\mathbf{F} \underset{\Omega}{\succ} \mathbf{F}'$$

with the condition (2) strengthened to

$$\sum c_i \omega_i - (s + k) \geq 0.$$

The system $\mathbf{F}$ shall be called Ω-high if for every k there is a $\mathbf{F}'$ such that (11) holds. Note the following lemma.

Lemma 1. *Suppose $\mathbf{F}$ is a Ω-bottomless system, then it is Ω-high.*

Proof. Fix a $k \geq 1$, then if $\mathbf{F}$ is Ω-bottomless there exists an infinite chain

$$\mathbf{F} = \mathbf{F}^{(1)} \underset{\Omega}{\succ} \mathbf{F}^{(2)} \underset{\Omega}{\succ} \ldots \underset{\Omega}{\succ} \mathbf{F}^{(k)} \underset{\Omega}{\succ} \cdots .$$

For each $m \geq 1$ we may write

$$\mathbf{F}^{(m+1)} = T_m \mathbf{F}_{\tau_m}^{(m)}$$

where $\tau_m \in GL(n, \mathcal{O}_K)$ with $v(\det \tau_m) = s_m$ and $T_m = (t_{ij,m})$ is an $(r \times r)$ upper triangular matrix with entries

$$t_{ij,m}(\mathbf{x}) = \pi^{-c_{i,m}} G_{ij,m}(\mathbf{x}),$$

for $1 \leq i \leq j \leq r$ where $c_{i,m} \geq 0$ and $G_{ij,m} \in \mathcal{O}_K[\mathbf{x}]$ is a form of degree[1]

$$\deg F_i - \deg F_j \geq 0$$

with $G_{ii} = 1$. Let Q be any positive integer such that $\omega_i \geq \frac{1}{Q}$ for every i, then by condition (2) we have that

$$\sum_i c_{i,m} \omega_i - (s_m + \tfrac{1}{Q}) \geq 0$$

[1] we assume once again that $d_1 \geq d_2 \geq \ldots \geq d_r$

for each $1 \leq m \leq r$. If T denotes the product of T_m and τ denotes the product of τ_m for $1 \leq m \leq kQ$, then we have that $\mathbf{F}^{(kQ+1)} = T\mathbf{F}_\tau$. Crucially we also have that

$$\sum_i \Big(\sum_m c_{i,m}\Big)\omega_i - \Big(\sum_m s_m + k\Big) \geq 0.$$

Therefore

$$\mathbf{F} \underset{\Omega}{\overset{k}{\succ}} \mathbf{F}^{(kQ+1)}$$

as required. □

We shall now note two Lemmata, the proofs of which can be found in Schmidt's paper [13, Lemmata 8 and 10].

Lemma 2. *Let $A_1, \ldots, A_l$ and $B_1, \ldots, B_m$ be linear forms with integer coefficients in the vector $\mathbf{x} = (x_1, \ldots, x_n)$. Let $\mathbf{x}_1, \mathbf{x}_2, \ldots$ be a sequence of vectors with*

$$A_i(\mathbf{x}_k) \geq 0 \quad (i = 1, \ldots, l;\ k = 1, 2, \ldots).$$

Then there exists a subsequence, $\mathbf{y}_1, \mathbf{y}_2, \ldots$ say, a constant B and an integer vector $\mathbf{a}$ with

$$A_i(\mathbf{a}) \geq 0 \quad (i = 1, \ldots, l)$$

such that

$$\lim_{k\to\infty} B_j(\mathbf{y}_k) = +\infty \quad \textit{for } j \textit{ with } B_j(\mathbf{a}) > 0$$

and

$$B_j(\mathbf{y}_k) \leq B \quad \textit{for } j \textit{ with } B_j(\mathbf{a}) \leq 0.$$

Before stating the next lemma we need to introduce some terminology. If $\mathbf{a}_1, \ldots, \mathbf{a}_s$ are linearly independent vectors in K^n, we call the set of linear combinations

$$c_1\mathbf{a}_1 + \cdots + c_s\mathbf{a}_s$$

a lattice, where $c_i \in \mathcal{O}_K$. We call $\mathbf{a}_1, \ldots, \mathbf{a}_s$, a basis of the lattice.

Lemma 3. *Suppose M is a sublattice of Λ. Then there exists a basis $\mathbf{l}_1, \ldots, \mathbf{l}_s$ of Λ and a basis of $\mathbf{m}_1, \ldots, \mathbf{m}_s$ of M such that*

$$\mathbf{m}_1 = \pi^{a_1}\mathbf{l}_1, \ldots, \mathbf{m}_s = \pi^{a_s}\mathbf{l}_s$$

for some non-negative integers $a_1, \ldots, a_s$.

Proof of Theorem 3*.* Suppose $\mathbf{F}$ is Ω-bottomless, then by Lemma 1 it is Ω-high. Hence for every $k \geq 1$ there are maps T_k and $\tau_k \in GL(n, \mathcal{O}_K)$ such that $T_k\mathbf{F}_{\tau_k}$ is defined over $\mathcal{O}_K$ and

$$\sum_i c_{i,k}\omega_i - (s_k + k) \geq 0 \tag{12}$$

where $v(\det \tau_k) = s_k$ and $T_k = (t_{ij,k})$ is an $(r \times r)$ upper triangular matrix with entries

$$t_{ij,k}(\mathbf{x}) = \pi^{-c_{i,k}} G_{ij,k}(\mathbf{x}),$$

for $1 \leq i \leq j \leq r$ where $c_{i,k} \geq 0$ and $G_{ij,k} \in \mathcal{O}_K[\mathbf{x}]$ is a form of degree

$$\deg F_i - \deg F_j \geq 0$$

with $G_{ii} = 1$. For ease of notation we write the ith row of the vector

$$diag(\pi^{c_{1,k}}, \ldots, \pi^{c_{r,k}}) T_k \mathbf{F}(\mathbf{x})$$

as

$$R_{i,k}(\mathbf{x}) := \sum_{j=i}^{r} G_{ij,k}(\mathbf{x}) F_j(\mathbf{x})$$

for $1 \leq i \leq r$. If Λ_k denotes the lattice $\tau_k \mathcal{O}_K^n$ then by assumption

$$\pi^{-c_{i,k}} R_{i,k}(\mathbf{x}) \in \mathcal{O}_K$$

for all $1 \leq i \leq r$ and every $\mathbf{x} \in \Lambda_k$. By Lemma 3, Λ_k has a basis

$$\Lambda_k : \pi^{u_1}\mathbf{u}_1, \ldots, \pi^{u_n}\mathbf{u}_n \tag{13}$$

where $\mathbf{u}_1, \ldots, \mathbf{u}_n$ is a basis of $\mathcal{O}_K^n$. Next we let $M_{R_{i,k}}$ denote the multilinear forms associated with $R_{i,k}$ for $1 \leq i \leq r$, then there exists some fixed non-negative integer γ such that

$$M_{R_{i,k}}(\mathbf{x}_1, \ldots, \mathbf{x}_{d_i}) \in \pi^{-\gamma}\mathcal{O}_K$$

for all $1 \leq i \leq r$ and any $\mathbf{x}_1, \ldots, \mathbf{x}_{d_i} \in \Lambda_k$. Consequently taking the basis vectors of Λ_k (13) we get

$$\pi^{-c_{i,k}} M_{R_{i,k}}(\pi^{u_{j_1}}\mathbf{u}_{j_1}, \ldots, \pi^{u_{j_{d_i}}}\mathbf{u}_{j_{d_i}}) \in \pi^{-\gamma}\mathcal{O}_K$$

or

$$|M_{R_{i,k}}(\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_{d_i}})| \leq p^{\gamma-(c_{i,k}-u_{j_1}-\ldots-u_{j_{d_i}})}$$

for all $1 \leq j_1, \ldots, j_{d_i} \leq n$.

Note that the $u_i, \mathbf{u}_i, c_{i,k}$ and $R_{i,k}$ all depend on k. Also note that since $\mathbf{u}_1, \ldots, \mathbf{u}_n$ is a basis of $\mathcal{O}_K^n$ we must have that $|\det(\mathbf{u}_1, \ldots, \mathbf{u}_n)| = 1$. By the compactness of $\mathcal{O}_K^n$ there must exist a subsequence of the sequence of integers $k = 1, 2, \ldots,$ such that on this subsequence $\mathbf{u}_1, \ldots, \mathbf{u}_n$ tend respectively to $\mathbf{a}_1, \ldots, \mathbf{a}_n$ and the forms $R_{1,k}, \ldots, R_{r,k}$ tend respectively to the forms $R_1, \ldots, R_r \in \mathcal{O}_K(\mathbf{x})$. To be clear the system $(R_1, \ldots, R_r)$ has a zero if and only if the system $\mathbf{F} = (F_1, \ldots, F_r)$ has a zero, since $G_{ii} = 1$ for each $1 \leq i \leq r$ in $R_{i,k}$. Moreover we have that $|\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)| = 1$, hence $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is a basis of $\mathcal{O}_K^n$. There exists a map σ defined over $\mathcal{O}_K$ such that

$$\sigma\mathbf{e}_i = \mathbf{a}_i, \quad \text{for } 1 \leq i \leq n.$$

With each k in our subsequence we define the vector

$$(A_1, \ldots, A_{n+r}) = (u_1, \ldots, u_n, c_{1,k}, \ldots, c_{r,k})$$

We also define

$$B = \sum_{i=1}^{r} c_{i,k}\omega_i - \sum_{j=1}^{n} u_j$$

and

$$B_i = c_{i,k} - (u_{j_1} + \cdots + u_{j_{d_i}})$$

for $1 \leq j_1, \ldots, j_{d_i} \leq n$ and $1 \leq i \leq r$. We apply Lemma 2 to the forms A_i, B, B_j, but before we do this note that the form B tends to $+\infty$. This is because $s_k = u_1 + \cdots + u_n$ and so by equation (12) one has

$$\sum_i c_{i,k}\omega_i - \sum_j u_j \geq k.$$

Now by Lemma 2 there is a vector

$$\mathbf{a} = (a_1, \ldots, a_n, b_1, \ldots, b_r)$$

with non-negative integer components, such that

$$\omega_1 b_1 + \cdots + \omega_r b_r - (a_1 + \cdots + a_n) > 0.$$

Moreover we have that B_i tends to $+\infty$ for all values i and $j_1, \ldots, j_{d_i}$ for which

$$b_i - (a_{j_1} + \cdots + a_{j_{d_i}}) > 0. \tag{14}$$

Taking the limit we obtain

$$M_{R_i}(\mathbf{a}_{j_1}, \ldots, \mathbf{a}_{j_{d_i}}) = 0$$

for all $j_1, \ldots, j_{d_i}$ satisfying (14). If we set $\mathbf{G} = \mathbf{R}_\sigma$ where $\mathbf{R} = (R_1, \ldots, R_r)$, then $\mathbf{G}$ satisfies precisely the conditions required to be Ω-special. Finally since $\mathbf{R}$ is equivalent to $\mathbf{F}$, we deduce that $\mathbf{F}$ is equivalent to a Ω-special system as required. □

4 Preliminaries for a Cubic and Quadratic form

Throughout this section and subsequent sections $\mathbf{F} = (F, G)$ will denote a system of a cubic and quadratic form in n variables defined over the ring of integers $\mathcal{O}_K$ of some $\mathfrak{p}$-adic field K. Let π denote a uniformizer for $\mathcal{O}_K$ and let $\mathbb{F}_q$ be the residue class field of K, where q denotes its cardinality.

Note the following corollary of Theorem 2.

Corollary 1. *Let (F, G) denote an arbitrary system of a cubic and quadratic form in $n \geq 14$ variables. Then there exists some $\omega_1 > 3$ and $\omega_2 > 2$ such that every (ω_1, ω_2)-bottomless system (F, G) has a non-trivial $\mathfrak{p}$-adic zero.*

Therefore it follows that to prove Theorem 1 it is sufficient to consider (ω_1, ω_2)-reduced systems for some $\omega_1 > 3$ and $\omega_2 > 2$.

Recall that we say that two systems $\mathbf{F}$ and $\mathbf{F}'$ are equivalent, if both systems are defined over $\mathcal{O}_K$ and

$$\mathbf{F}' = T\mathbf{F}_\tau$$

where

$$T = \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix},$$

for some linear form $L \in \mathcal{O}_K[\mathbf{x}]$ and $\tau \in GL(n, \mathcal{O}_K)$ with $v(\det \tau) = 0$. The

order $o(\mathbf{F})$, of a system $\mathbf{F}$ is the least positive integer m such that $\mathbf{F}$ is equivalent to a system that contains m variables explicitly. Also recall the h-invariant of $\mathbf{F}$, denoted $h(\mathbf{F})$ is the least integer h such that we can write

$$F(\mathbf{x}) = x_1 H_{11}(\mathbf{x}) + \ldots + x_h H_{1h}(\mathbf{x}) \pmod{\pi},$$

and

$$G(\mathbf{x}) = x_1 H_{21}(\mathbf{x}) + \ldots + x_h H_{2h}(\mathbf{x}) \pmod{\pi},$$

for all systems equivalent to $\mathbf{F}$. We similarly define the h-invariant of a single form in the obvious way. Note here that since $\mathbf{F}$ is defined over $\mathcal{O}_K$, considering it modulo π is well defined. Note the following lemma which will play a crucial part in our proof.

Lemma 4. *Suppose $\mathbf{F} = (F, G)$ is (α, β)-reduced, for some $\alpha > 3$ and $\beta > 2$, then*

$$h(G) > 2, \quad h(F - LG) > 3 \quad \textit{and} \quad h(\mathbf{F}) > 5$$

for every linear form $L(\mathbf{x}) \in \mathcal{O}_K[\mathbf{x}]$.

Proof. Let τ_r be the diagonal $n \times n$ matrix which has π as its first r entries and 1 otherwise. If $h(G) \leq 2$ then we may write

$$G = x_1 L_1 + x_2 L_2 \pmod{\pi},$$

for some linear forms $L_i \in \mathcal{O}_K[\mathbf{x}]$. If we let

$$T = \begin{pmatrix} 1 & 0 \\ 0 & \pi^{-1} \end{pmatrix},$$

then $\mathbf{F}' = T\mathbf{F}_{\tau_2}$ is defined over $\mathcal{O}_K$. However $\beta - 2 > 0$, contradicting our assumption that $\mathbf{F}$ is (α, β)-reduced (cf. condition (2), p.2). If $h(F - LG) \leq 3$ for some linear form $L(\mathbf{x}) \in \mathcal{O}_K[\mathbf{x}]$, then we may write

$$F - LG = x_1 Q_1 + x_2 Q_2 + x_3 Q_3 \pmod{\pi},$$

for some quadratic forms $Q_i \in \mathcal{O}_K[\mathbf{x}]$. If we let

$$T = \begin{pmatrix} \pi^{-1} & -\pi^{-1} L \\ 0 & 1 \end{pmatrix},$$

then $\mathbf{F}' = T\mathbf{F}_{\tau_3}$ is defined over $\mathcal{O}_K$. However $\alpha - 3 > 0$, contradicting that $\mathbf{F}$ is (α, β)-reduced. Finally if $h(\mathbf{F}) \leq 5$ then we may write

$$\mathbf{F}' = (F - LG, G) = (x_1 Q_1 + \ldots + x_5 Q_5, x_1 L_1 + \ldots + x_5 L_5) \pmod{\pi},$$

for some quadratic and linear forms $Q_i, L_i, L \in \mathcal{O}_K[\mathbf{x}]$. This time we let

$$T = \begin{pmatrix} \pi^{-1} & 0 \\ 0 & \pi^{-1} \end{pmatrix},$$

then $\mathbf{F}'' = T\mathbf{F}'_{\tau_5}$ is defined over $\mathcal{O}_K$. However $\alpha + \beta - 5 > 0$, contradicting that $\mathbf{F}$ is (α, β)-reduced. $\square$

5 Reduced systems

In this section we will work with our system $\mathbf{F} = (F, G)$ modulo π, which from now on we shall denote as $\mathbf{f} = (f, g)$. We assume that $\mathbf{F}$ is (α, β)-reduced, for some $\alpha > 3$ and $\beta > 2$. Hence $\mathbf{f}$ will satisfy the conclusion of Lemma 4 viz.

$$h(g) > 2 \quad h(f - lg) > 3 \quad \text{and} \quad h(\mathbf{f}) > 5 \tag{15}$$

for every linear form $l(\mathbf{x}) \in \mathbb{F}_q[\mathbf{x}]$. We also denote $m = o(\mathbf{f})$.

The aim of this section is to show that we can find a non singular zero of $\mathbf{f}$ which by Hensel's Lemma will lift to give us a zero of our original system $\mathbf{F}$. For clarity we outline the steps we will take in order to prove this:

Step 1: We prove that we can find a zero $\mathbf{e}_1$ say, of $\mathbf{f}$ such that $\nabla g \neq 0$. Therefore we are able to write our system $\mathbf{f}$ in the shape

$$\begin{aligned} f(\mathbf{x}) &= x_1^2 f_1 + x_1 f_2 + f_3 \\ g(\mathbf{x}) &= x_1 g_1 + g_2 \end{aligned}$$

where $g_1 \not\equiv 0$. If $\mathbf{e}_1$ is a non singular zero then we're done. Otherwise we can find some $\lambda \in \mathbb{F}_q$ such that $f_1 = \lambda g_1$. We now consider the equivalent system:

$$\begin{aligned} (f - \lambda x_1 g)(\mathbf{x}) &= x_1(f_2 - \lambda g_2) + f_3 \\ g(\mathbf{x}) &= x_1 g_1 + g_2. \end{aligned}$$

We may therefore assume that $\mathbf{f}$ is equivalent to one of two situations:

(i) $\deg_{x_1} f = 0$,

(ii) $\deg_{x_1} f = 1$.

Step 2: In case (i) we show that we may write $\mathbf{f}$ as

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_2) \\ g(\mathbf{x}) &= x_1 x_2 + g_2(\mathbf{x}_3) \end{aligned}$$

where we define $\mathbf{x}_i := (x_i, x_{i+1}, \ldots, x_m)$. Next we will find a non singular zero $\mathbf{x}_2$ of f such that $x_2 \neq 0$. Hence by setting $x_1 = x_2^{-1} g_3(\mathbf{x}_3)$, we get a non singular zero of the system $\mathbf{f}$ as required.

Step 3: In case (ii) we show that we can write $\mathbf{f}$ as

$$\begin{aligned} f(\mathbf{x}) &= x_1 f_2(\mathbf{x}_3) + f_3(\mathbf{x}_2) \\ g(\mathbf{x}) &= x_1 x_2 + g_2(\mathbf{x}_3). \end{aligned}$$

Next we define the quartic form

$$H(\mathbf{x}_2) := x_2 f_3(\mathbf{x}_2) - (f_2 g_2)(\mathbf{x}_3).$$

It will follow that if we can find a non singular zero of H such that $x_2 \neq 0$ then we can find a non singular zero of the system $\mathbf{f}$. Finding a non singular zero of H such that $x_2 \neq 0$ requires a blend of ideas which utilizes the information (15) we have at our disposal about the h-invariant of the system.

Having described the outline of the proof we proceed with Step 1.

Step 1

Throughout this step and subsequent steps we will need to make use of three important Lemmata, the first being attributed to Warning [16] (for example see [12, Theorem 1E, p.137]).

Lemma 5. *Let $F_1, \ldots, F_r$ be a system of forms of degrees $d_1, \ldots, d_r$ respectively in m variables over $\mathbb{F}_q$. If $m > \delta = \sum d_i$, then the system $F_1, \ldots, F_r$ has at least $q^{m-\delta}$ common affine $\mathbb{F}_q$-rational zeros.*

The next Lemma comes from a book of Schmidt [12, Lemma 3A, p.147].

Lemma 6. *Let F be a non zero polynomial over $\mathbb{F}_q$ in m variables of total degree d. Then the number N_a of affine zeros of F in $\mathbb{F}_q^n$ satisfies*

$$N_a \leq dq^{m-1}.$$

The final lemma in our toolbox is due to Leep & Yeomans [9].

Lemma 7. *Let $P \in \mathbb{F}_q[x, y]$ be an absolutely irreducible polynomial of degree d. Then the number N of non-singular zeros of P satisfies*

$$N \geq q + 1 - \frac{1}{2}(d-1)(d-2)[2\sqrt{q}],$$

where $[\gamma]$ denotes the least integer not exceeding γ.

Proof. Write S to denote the number of $\mathbb{F}_q$ singular zeros of P. Then if the curve defined by $P(x, y) = 0$ has genus g (not to be confused with the quadratic form g), it follows from Corollary 1 of Leep & Yeomans [9] that

$$|N + S - (q+1)| \leq g([2\sqrt{q}] - 1) + \frac{1}{2}(d-1)(d-2).$$

Next we use the above estimate together with the following bound on the genus

$$g \leq \frac{1}{2}(d-1)(d-2) - S,$$

which comes from Lemma 1 of [9], to obtain the required bound

$$N \geq q + 1 - \frac{1}{2}(d-1)(d-2)[2\sqrt{q}].$$

□

We say that the system $\mathbf{f}$ is equivalent the system $\mathbf{f}' = (f', g')$, if

$$\begin{aligned} f'(\mathbf{x}) &= (f - lg)(\tau\mathbf{x}) \\ g'(\mathbf{x}) &= g(\tau\mathbf{x}) \end{aligned}$$

for some linear form $l \in \mathbb{F}_q[\mathbf{x}]$ and $\tau \in GL(n, \mathbb{F}_q)$. That being said we let $a \geq 0$ denote the maximum integer such that we can write

$$\begin{aligned} f(\mathbf{x}) &= f(y_1, \ldots, y_a, z_1, \ldots, z_b) \\ g(\mathbf{x}) &= g(z_1, \ldots, z_b) \end{aligned}$$

where $o(\mathbf{f}) = m = a + b$ and $o(g) = b$, amongst all systems equivalent to $\mathbf{f} = (f, g)$. We introduce the notation $\mathbf{y} = (y_1, \ldots, y_a)$ and $\mathbf{z} = (z_1, \ldots, z_b)$. Throughout this step we shall assume that every zero of $\mathbf{f}$ is such that $\mathbf{z} = \mathbf{0}$, otherwise the zero will be such that $\nabla g \neq \mathbf{0}$, which completes Step 1. Since $m \geq h(\mathbf{f}) \geq 6$, then by Lemma 5 there exists a zero $\mathbf{e}_1$ say of $\mathbf{f}$. Also note that since $\mathbf{f}(x_1, 0, \ldots, 0) = 0$ for all $x_1 \in \mathbb{F}_q$, then $m-1 \geq h(\mathbf{f}) \geq 6$. Let $N(\mathbf{f})$ denote the number of affine zeros of the system $\mathbf{f}$ over $\mathbb{F}_q$. Hence by Lemma 5,

$$N(\mathbf{f}) \geq q^2.$$

So there exists another zero of $\mathbf{f}$ not in the affine span of $\mathbf{e}_1$, say $\mathbf{e}_2$. If all the zeros of $\mathbf{f}$ are in the affine span of $\{\mathbf{e}_1, \mathbf{e}_2\}$ and $q > 2$ then $m - 2 \geq h(\mathbf{f}) \geq 6$. Hence by Lemma 5,

$$N(\mathbf{f}) \geq q^3.$$

Otherwise there is a zero of $\mathbf{f}$ not in the span of $\{\mathbf{e}_1, \mathbf{e}_2\}$. In either case we may assume that there are at least 3 linearly independent zeros of $\mathbf{f}$, say $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$.

If any of these zeros are such that $\nabla g \neq \mathbf{0}$ then we have completed Step 1. Otherwise we may assume each of these zeros $\mathbf{e}_i$ are singular for g, which implies that $a \geq 3$. Note that we may write

$$f(\mathbf{x}) = f_a(\mathbf{y}) + f_{a,b}(\mathbf{y}, \mathbf{z})$$

where $f_a(\mathbf{y}) := f(\mathbf{y}, \mathbf{0})$.

Suppose that $b \geq 5$, then by Lemma 5

$$N(\mathbf{f}) \geq q^{a+b-5} \geq q^a.$$

If all the zeros of $\mathbf{f}$ are such that $\mathbf{z} = \mathbf{0}$, then by the above inequality $\mathbf{f}$ must vanish in the span of $\{\mathbf{e}_1, \ldots, \mathbf{e}_a\}$. Hence we have $m - a \geq h(\mathbf{f}) \geq 6$, so that

$$N(\mathbf{f}) \geq q^{a+1}$$

Therefore we conclude that in any case we can find a zero of $\mathbf{f}$ such that $\mathbf{z} \neq \mathbf{0}$, completing Step 1 provided $b \geq 5$. Moreover since $b \geq h(g) \geq 3$ we can find a zero $\mathbf{e}_{a+1}$ say of g, therefore as before we have that $b - 1 \geq h(g) \geq 3$. Hence we may assume from now on that $b = 4$, otherwise as the above argument shows $b \geq 5$ would allow us to find a zero with $\mathbf{z} \neq \mathbf{0}$.

Next we would like to find a zero $\mathbf{e}_1$ say of f_a, such that the variable x_1 appears in f_a. Let $a' = o(f_a)$ and write $f_a(\mathbf{y}) = f_a(\mathbf{y}')$, where $\mathbf{y}' = (y_1, \ldots, y_{a'})$ after a non singular change of the variables $\mathbf{y}$. Note the following relationship

$$6 \leq h(\mathbf{f}) \leq h(f_a) + b = h(f_a) + 4.$$

Therefore $a' \geq h(f_a) \geq 2$. We now show that we can find a zero of $\mathbf{f}$ such that $\mathbf{y}' \neq \mathbf{0}$. Suppose all zeros of $\mathbf{f}$ were such that $\mathbf{y}' = \mathbf{0}$ and $\mathbf{z} = \mathbf{0}$, that leaves $a - a'$ variables which are non zero for each solution of $\mathbf{f} = \mathbf{0}$. However

$$N(\mathbf{f}) \geq q^{a'+(a-a')+b-5} \geq q^{a-a'+1}.$$

Which implies that either we can find a zero of $\mathbf{f}$ such that $\mathbf{z} \neq \mathbf{0}$ (completing Step 1) or we can find a zero such that $\mathbf{y}' \neq \mathbf{0}$. Hence we may assume that $\mathbf{e}_1$ is a non trivial zero of f_a. So if $\mathbf{e}_{a+1}$ is a zero of g we can assume that $f(\mathbf{e}_{a+1}) \neq 0$

otherwise we would have found the required zero to complete this step. So we have that

$$f(\mathbf{x}) = y_1 f_2(y_1, \ldots, y_a) + f_3(y_2, \ldots, y_a) + f_{a,b}(\mathbf{y}, \mathbf{z}),$$

where $f_2 \not\equiv 0$ and $f_{a,b}(\mathbf{e}_{a+1}) \neq 0$. By Lemma 6 for $q > 2$, we can find a vector $\mathbf{e}_2$ say such that $f_2(\mathbf{e}_2) \neq 0$. Finally we consider the following slice of the cubic form.

$$S(X, Y, Z) := f(X\mathbf{e}_1 + Y\mathbf{e}_2 + Z\mathbf{e}_{a+1}) = X^2 u(Y, Z) + Xv(Y, Z) + w(Y, Z), \quad (16)$$

where $v(Y, 0) \neq 0$ and $w(0, Z) \neq 0$. Note that if we can find a zero of S such that $Z \neq 0$, then we can have found a zero of $\mathbf{f}$ with $\mathbf{z} \neq \mathbf{0}$ which completes Step 1. Note the following Lemma.

Lemma 8. *$S(X, Y, Z)$ has a zero with $Z \neq 0$.*

Proof. If S is absolutely irreducible we can set $Z = 1$ and apply Lemma 7, to deduce the existence of a zero over $\mathbb{F}_q$ for all q. If S is reducible then we may assume that either it is the product of 3 conjugate linear factors or the product of a linear factor defined over $\mathbb{F}_q$ and a quadratic factor.

It cannot be the product of 3 conjugate linear factors otherwise a X^3 term would appear in S. Hence we may assume that S is the product of a linear factor over $\mathbb{F}_q$ and a quadratic factor. Note that Z does not divide S, since $v(Y, 0) \neq 0$. Therefore we can choose (X, Y, Z) so that we set the linear factor equal to 0 and have $Z \neq 0$, completing the proof. □

This completes Step 1 viz. we can find a vector $\mathbf{e}_1$ such that $\mathbf{f} = \mathbf{0}$ and $\nabla g \neq \mathbf{0}$.

Step 2

In this step we may assume that our zero $\mathbf{e}_1$ of $\mathbf{f}$ is such that $\nabla g \neq \mathbf{0}$ and $\deg_{x_1} f = 0$. Therefore we may immediately write

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_2) \\ g(\mathbf{x}) &= x_1 g_1(\mathbf{x}_2) + g_2(\mathbf{x}_2), \end{aligned}$$

where $g_1 \not\equiv 0$. By a non singular change of variables we can assume that $g_1 = x_2$. Hence we may write

$$\begin{aligned} g(\mathbf{x}) &= x_1 x_2 + g_2(\mathbf{x}_2) \\ &= x_2(x_1 + \lambda x_2 + L(\mathbf{x}_3)) + \hat{g}_2(\mathbf{x}_3), \end{aligned}$$

for some constant $\lambda \in \mathbb{F}_q$ and linear form $L(\mathbf{x}_3)$. So by mapping x_1 to $x_1 - \lambda x_2 - L(\mathbf{x}_3)$ and writing g_2 to denote $\hat{g}_2$ we have

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_2) \\ g(\mathbf{x}) &= x_1 x_2 + g_2(\mathbf{x}_3). \end{aligned}$$

Our goal now is to find a non-singular zero of $f(\mathbf{x}_2)$ such that $x_2 \neq 0$. Then by setting $x_1 = -x_2^{-1} g_2(\mathbf{x}_3)$, we obtain a non-singular zero of our system $\mathbf{f}$, as required. We can appeal to a result of Lewis & Schuur [11, Theorem 3] to immediately answer this exact question viz.

Theorem 4 (Lewis & Schuur, 1973)**.** *Let k be a finite field of cardinality $q \geq 5$. Let F be a non-degenerate cubic form over k such that $o(F) \geq 4$ and let L be a linear form over k. Then F has a k-point which is a non-singular zero of F and is not a zero of L.*

To keep with the slicing theme we will prove our own version of Theorem 4 by employing the information we have about the h-invariant. First we need a lemma.

Lemma 9. *Let $F(x_1, \ldots, x_m)$ be a non-degenerate cubic form over any finite field, such that $o(F) = m \geq 4$. Then F has a non-singular zero.*

Proof. By Lemma 5, F has a non-trivial zero $\mathbf{e}_1$ say. Therefore we may write

$$F(\mathbf{x}) = x_1^2 F_1(\mathbf{x}_2) + x_1 F_2(\mathbf{x}_2) + F_3(\mathbf{x}_2).$$

If $F_1 \not\equiv 0$ then $\mathbf{e}_1$ is a non-singular zero. Otherwise,

$$F(\mathbf{x}) = x_1 F_2(\mathbf{x}_2) + F_3(\mathbf{x}_2)$$

where $F_2 \not\equiv 0$ since F is non-degenerate. In any finite field we can find a point $\mathbf{x}_2 \in \mathbb{F}_q^{m-1}$ such that $F_2(\mathbf{x}_2) \neq 0$. Therefore we obtain a non-singular zero of F by setting $x_1 = (-F_2^{-1} F_3)(\mathbf{x}_2)$, as required. $\square$

Next we need to consider $\delta := \deg_{x_2} f(\mathbf{x}_2)$. If $\delta = 0$, then by Lemma 9 we can find a non-singular zero of f and set $x_2 = 1$ and $x_1 = -g_2$ to obtain non-singular zero of $\mathbf{f}$. If $\delta = 1$ then

$$f(\mathbf{x}_2) = x_2 f_2(\mathbf{x}_3) + f_3(\mathbf{x}_3),$$

where $f_2 \not\equiv 0$ since $\delta = 1$ and $f_3 \not\equiv 0$ since $h(f) > 3$. By Lemma 6 we can find a vector $\mathbf{x}_3 \in \mathbb{F}_q^{n-2}$ such that $f_2(\mathbf{x}_3), f_3(\mathbf{x}_3) \neq 0$, provided $q > 5$. Therefore by setting $x_2 = (-f_2^{-1} f_3)(\mathbf{x}_3) \neq 0$ and $x_1 = -x_2^{-1} g_2(\mathbf{x}_3)$, we get a non-singular zero of $\mathbf{f}$ as required. If $\delta = 2$, then $\mathbf{e}_2$ is a non-singular zero of $\mathbf{f}$ as required. Therefore we may assume that $\delta = 3$.

By Lemma 9 we can find a non-singular zero $\mathbf{a} = (a_2, \ldots, a_n)$ say of $f(\mathbf{x}_2)$. If $a_2 \neq 0$, then we can find a non-singular zero of $\mathbf{f}$ by setting $x_1 = -a_2^{-1} g_2(a_3, \ldots, a_n)$. So if $a_2 = 0$, we can make a change of variables so that $\mathbf{a} = \mathbf{e}_3$. Therefore

$$f(\mathbf{x}_2) = x_3^2 f_1'(x_2, \mathbf{x}_4) + x_3 f_2'(x_2, \mathbf{x}_4) + f_3'(x_2, \mathbf{x}_4) + f_3''(\mathbf{x}_4),$$

where $\deg_{x_2} f_3' = 3$ (since $\delta = 3$), $f_1' \not\equiv 0$ (since $\mathbf{e}_3$ is a non-singular zero) and $f_3'' \not\equiv 0$ (since $h(f) > 3$). By Lemma 6 we can find a vector $\mathbf{e}_4$ such that, $f_2'(x_2, \mathbf{e}_4), f_3''(\mathbf{e}_4) \neq 0$, provided $q > 5$. We now consider the slice

$$\begin{aligned} T(X, Y, Z) &= f(X\mathbf{e}_2 + Y\mathbf{e}_3 + Z\mathbf{e}_4) \\ &= cX^3 + Y^2 u(X, Z) + Y v(X, Z) + w(X, Z) + dZ^3, \end{aligned}$$

where $c, d \neq 0$, $u(X, 1) \neq 0$ and $\deg_X w \leq 2$.

Lemma 10. *$T(X, Y, Z)$ has a non-singular zero such that $X \neq 0$, provided $q > 3$.*

Proof. If T is absolutely irreducible we can set $X = 1$ and apply Lemma 7 to deduce the existence of a zero over $\mathbb{F}_q$ for all q. If T is reducible then either it is the product of 3 conjugate linear factors or the product of a linear factor defined over $\mathbb{F}_q$ and a quadratic factor.

It cannot be the product of 3 conjugate linear factors otherwise a Y^3 term would appear in T. Hence we may assume that T is the product of a linear factor over $\mathbb{F}_q$ and a quadratic factor. Note that X does not divide T, since $d \neq 0$. Therefore we can assume that the linear factor is $X - l(Y, Z)$, where $l \not\equiv 0$. Let $Q(X, Y, Z)$ denote the quadratic factor of T. If $X - l(Y, Z)$ does not divide Q then $Q(l(Y, Z), Y, Z) \not\equiv 0$. By Lemma 6 provided $q > 3$ we can find some (Y, Z) such that $l(Y, Z), Q(l(Y, Z), Y, Z) \neq 0$. Therefore by setting $X = l(X, Y) \neq 0$ we get the required non-singular zero. On the other hand if $X - l(Y, Z)$ does divide Q then we may write

$$T(X, Y, Z) = (X - l(Y, Z))^2(cX - l'(Y, Z)).$$

Now note that since $\deg_Y T = 2$, $cX - l'(Y, Z)$ cannot divide $X - l(Y, Z)$. So by letting $Q(X, Y, Z) = (X - l(Y, Z))^2$ we may find the required non-singular zero as before. $\square$

This completes Step 2 viz. If $\mathbf{e}_1$ is a zero of $\mathbf{f}$ such that $\nabla g \neq \mathbf{0}$ and $\deg_{x_1} f = 0$, then we can find a non-singular zero of $\mathbf{f}$ provided $q > 5$.

Step 3

In this step we shall assume that $\mathbf{e}_1$ is a zero of $\mathbf{f}$ such that $\nabla g \neq \mathbf{0}$ and $\deg_{x_1} f = 1$. Therefore we can write

$$\begin{aligned} f(\mathbf{x}) &= x_1 f_2(\mathbf{x}_2) + f_3(\mathbf{x}_2) \\ g(\mathbf{x}) &= x_1 g_1(\mathbf{x}_2) + g_2(\mathbf{x}_2) \end{aligned}$$

Recall (cf. beginning of Step 2) that we can make a change of basis so that

$$g(\mathbf{x}) = x_1 x_2 + g_2(\mathbf{x}_3),$$

and $\deg_{x_1} f = 1$. Moreover by subtracting linear multiples of g from f we may assume that $f_2 = f_2(\mathbf{x}_3)$. As mentioned in the outline of this step, we define

$$H(\mathbf{x}_2) = x_2 f_3(\mathbf{x}_2) - (f_2 g_2)(\mathbf{x}_3).$$

Our strategy is to find a vector $\mathbf{x}_2$ to be able to apply the following lemma.

Lemma 11. *If we can find a non-singular zero $\mathbf{x}_2 \in \mathbb{F}_q^{n-1}$ say of H such that $x_2 \neq 0$, then $(-x_2^{-1} g_2(\mathbf{x}_3), \mathbf{x}_2) \in \mathbb{F}_q^n$ is a non-singular zero of $\mathbf{f}$.*

Proof. First we show that $(-x_2^{-1} g_2(\mathbf{x}_3), \mathbf{x}_2)$ is a zero of $\mathbf{f}$. It is clear that $\mathbf{x}$ is a zero of g since, $x_1 = -x_2^{-1} g_2(\mathbf{x}_3)$. Also since $x_2 \neq 0$ we have

$$\begin{aligned} H(\mathbf{x}_2) &= x_2(f_3(\mathbf{x}_3) - x_2^{-1}(g_2 f_2)(\mathbf{x}_3)) \\ &= x_2 f(\mathbf{x}). \end{aligned}$$

Therefore $f(\mathbf{x}) = 0$. Next suppose $\mathbf{x}$ is a singular zero of $\mathbf{f}$, then the following gradient vectors must be linearly dependent

$$\begin{aligned} \nabla f(\mathbf{x}) &= (f_2, f_{32}, x_1 f_{2i} + f_{3i}) \\ \nabla g(\mathbf{x}) &= (x_2, x_1, g_{2i}) \end{aligned}$$

where f_{2i}, f_{3i}, g_{2i} denotes $\frac{\partial f_2}{\partial x_i}, \frac{\partial f_3}{\partial x_i}, \frac{\partial g_2}{\partial x_i}$ respectively for $2 \leq i \leq n$. Consequently we have the following vector identity

$$f_2 \nabla g = x_2 \nabla f.$$

Looking at the components of this identity we have that

$$x_1 f_2 = x_2 f_{32} \tag{17}$$

$$f_2 g_{2i} = x_2(x_1 f_{2i} + f_{3i}) \tag{18}$$

for $3 \leq i \leq n$. Next we note the gradient vector of H,

$$\nabla H(\mathbf{x}_2) = (f_3 + x_2 f_{32}, x_2 f_{3i} - g_{2i} f_2 - g_2 f_{2i}).$$

Let $(\nabla H)_j$ denote the jth component of the vector ∇H for $1 \leq j \leq n-1$. Then by (17),

$$(\nabla H)_1 = f_3 + x_2 f_{32} = f_3 + x_1 f_2 = f = 0.$$

Also by (18) for $2 \leq j \leq n-1$ we have,

$$\begin{aligned} (\nabla H)_j &= x_2 f_{3i} - g_{2i} f_2 - g_2 f_{2i} \\ &= x_2 f_{3i} - g_2 f_{2i} - x_2(x_1 f_{2i} + f_{3i}) \\ &= -f_{2i}(x_1 x_2 + g_2) = -f_{2i} g = 0. \end{aligned}$$

Therefore $\nabla H = \mathbf{0}$, a contradiction. Hence $\mathbf{x}$ is a non-singular zero of $\mathbf{f}$ as required. $\square$

We shall now show that we can find a non-singular zero $\mathbf{x}_2$ say of H such that $x_2 \neq 0$. Of course from the outset it may be possible that H is the product of a quadratic form Q say with itself. Then if $H = 0$ we must have $Q = 0$, therefore

$$\frac{\partial H}{\partial x_i} = 2Q \frac{\partial Q}{\partial x_i} = 0, \quad \text{for all } \ 2 \leq i \leq n$$

implying that every zero of H is singular. We shall show that this cannot happen, more precisely we will prove the following.

Lemma 12. *Suppose the h-invariant condition* (15) *viz.*

$$h(g) > 2 \quad h(f - lg) > 3 \quad \textit{and} \quad h(\mathbf{f}) > 5$$

then the form

$$H(\mathbf{x}_2) = x_2 f_3(\mathbf{x}_2) - (g_2 f_2)(\mathbf{x}_3)$$

is necessarily absolutely irreducible over $\mathbb{F}_q$.

Proof. As an outline we will distinguish between the cases in which H is either the product of two quadratic forms or the product of a linear form and absolutely irreducible cubic form. Suppose H factors over $\bar{\mathbb{F}}_q$, therefore we can write

$$\begin{aligned} H(\mathbf{x}_2) &= x_2 f_3(\mathbf{x}_2) - (g_2 f_2)(\mathbf{x}_3) \\ &= (A(\mathbf{x}_3) + x_2 B(\mathbf{x}_2))(A'(\mathbf{x}_3) + x_2 B'(\mathbf{x}_2)) \\ &= AA' + x_2(AB' + A'B) + x_2^2 BB' \end{aligned}$$

for some forms $A, A', B, B' \in \bar{\mathbb{F}}_q[\mathbf{x}_2]$. Hence

$$-g_2 f_2 = AA'. \tag{19}$$

Also by looking at the x_2 coefficient we deduce

$$f_3 = AB' + A'B + x_2 BB'. \tag{20}$$

Case (a): Suppose H is the product of a linear and absolutely irreducible cubic factor over some extension K say of $\mathbb{F}_q$. Then by considering the action of $\mathrm{Gal}(K : \mathbb{F}_q)$ on the factors, it is easy to see that each factor must be defined over $\mathbb{F}_q$. Suppose $(A + x_2 B)$ is the linear factor, then (19) implies that A divides either g_2 or f_2. It cannot divide g_2 since if it did then by setting the two linear forms $A, x_1 = 0$ we would have that $g = 0$. Hence $h(g) \leq 2$, contradicting (15). So $f_2 = AL$ for some linear form L defined over $\mathbb{F}_q$. Therefore $A' = -g_2 L$. So by (20) we have

$$\begin{aligned} f(\mathbf{x}) &= x_1 f_2 + f_3 \\ &= x_1 AL + AB' - g_2 LB + x_2 BB'. \end{aligned}$$

Therefore by setting the linear forms $A, L, x_2 = 0$ we deduce that $f = 0$. Hence $h(f) \leq 3$, contradicting (15).

Case (b): Suppose $A = g_2$ and $A' = -f_2$. By (20) we can write

$$\begin{aligned} f(\mathbf{x}) &= x_1 f_2 + f_3 \\ &= -x_1 A' + AB' + A'B + x_2 BB' \\ &= B'g + (-x_1 + B)(A' + x_2 B') \end{aligned}$$

recalling that

$$\begin{aligned} g(\mathbf{x}) &= x_1 x_2 + g_2 \\ &= x_1 x_2 + A. \end{aligned}$$

Since $A = g_2$ and $A' = -f_2$ are defined over $\mathbb{F}_q$, then either the factors $(A + x_2 B)$ and $(A' + x_2 B')$ are defined over $\mathbb{F}_q$ or the quadratic extension of $\mathbb{F}_q$. In the former case B and B' have coefficients in $\mathbb{F}_q$ and by setting $x_1, B, B' = 0$ we deduce that $h(f) \leq 3$, a contradiction to (15). In the latter case we may assume that B and B' are conjugates of each other over the quadratic extension of $\mathbb{F}_q$. Hence we may write

$$B(\mathbf{x}_2) = l_1(\mathbf{x}_2) + \alpha l_2(\mathbf{x}_2)$$

and

$$B(\mathbf{x}_2) = \alpha l_1(\mathbf{x}_2) + l_2(\mathbf{x}_2)$$

for some linear forms $l_1, l_2 \in \mathbb{F}_q[\mathbf{x}_2]$ and $\alpha \in \bar{\mathbb{F}}_q$. So if we set $l_1, l_2, x_1 = 0$ we deduce that $h(f) \leq 3$, a contradiction as before. This completes Case (b).

Before moving onto the final case we shall make a few remarks. If $\mathrm{rank}(g_2) \geq 3$ and H is the product of two quadratic factors then we can assume that we're in Case (b). This is because the rank condition implies that g_2 is absolutely irreducible and therefore the condition (19) forces $\{A, A'\} = \{f_2, -g_2\}$, up to a scalar multiple in $\mathbb{F}_q$. We are therefore left to deal with the case in which $\mathrm{rank}(g_2) \leq 2$.

Case (c): Suppose $\mathrm{rank}(g_2) \leq 2$. So g_2 is reducible over $\bar{\mathbb{F}}_q$ and we may write

$$g_2(\mathbf{x}_3) = g_2(x_3, x_4) = l_1(x_3, x_4) l_2(x_3, x_4)$$

for some linear forms l_1, l_2 defined over $\bar{\mathbb{F}}_q$. We can assume that H is the product of two quadratic factors since the other possibility is dealt with in Case (a). Either $\{A, A'\} = \{f_2, -g_2\}$, up to a scalar multiple in which instance we are in Case (b), or we can assume that l_1 divides A and l_2 divides A'. Hence $x_3, x_4 = 0$ implies that $A, A' = 0$. So as before we write

$$f(\mathbf{x}) = x_1 f_2 + AB' + A'B + x_2 BB' \quad (21)$$

$$g(\mathbf{x}) = x_1 x_2 + l_1 l_2. \quad (22)$$

So by setting the linear forms $x_1, x_2, x_3, x_4 = 0$ we deduce that $\mathbf{f} = \mathbf{0}$. Hence $h(\mathbf{f}) \leq 4$, a contradiction to (15).

This completes the proof of the lemma. ☐

We are now in a position to find a non-singular zero $\mathbf{x}_2$ for H where $x_2 \neq 0$ which by Lemma 11 will imply that there is a non-singular zero of $\mathbf{F}$. To do this we employ a slicing approach, following an idea used by Wooley [17] for the case of degree 7 and 11 forms. We do this owing to the sharp bounds that are available for point counting on curves over $\mathbb{F}_q$ opposed to hypersurfaces.

Before stating the next Lemma, we shall need to introduce some notation. Let L be a field and consider a polynomial $f \in L[x_0, x_1, \ldots, x_n]$. When $\xi \in L^{3n+1}$, we write $f|_\xi = f|_\xi(X, Y)$ to denote the sliced polynomial

$$f(\xi_0 + X, \xi_1 + \xi_{n+1}X + \xi_{2n+1}Y, \ldots, \xi_n + \xi_{2n}X + \xi_{3n}Y).$$

Next we shall note the following result of Cafure & Matera [5].

Lemma 13. *Let $f \in \mathbb{F}_q[x_0, \ldots, x_n]$ be an absolutely irreducible polynomial of degree $d \geq 2$. Then the number of slices $\xi \in \mathbb{F}_q^{3n+1}$, for which the polynomial $f|_\xi$ is not absolutely irreducible, is at most $\frac{1}{2}(3d^4 - 4d^3 + 5d^2)q^{3n}$.*

Proof. This is Corollary 3.2 of [5]. ☐

Finally by Lemma 12 and Lemma 13 if $q \geq 296$ there exists a slice $H|_\xi$ of

$$H(\mathbf{x}) = x_2 f_3(\mathbf{x}_2) - (g_2 f_2)(\mathbf{x}_3)$$

which is an absolutely irreducible curve. Moreover on this slice we must have that the x_2 component is not identically zero otherwise $H|_\xi$ would factor into the product $(g_2|_\xi)(f_2|_\xi)$. Finally by Lemma 7 (p.11), taking $q > 293$ is more than sufficient to ensure the existence of a non-singular zero of $H|_\xi$ for which $x_2 \neq 0$ and so of H as required. This completes the proof of Theorem 1

References

[1] E. Artin. *The collected papers of Emil Artin.* Addison-Wesley, London, 1965.

[2] J. Ax and S. Kochen. Diophantine problems over local fields. I. *Amer. J. Math.*, 87:605–630, 1965.

[3] B. J. Birch, D. J. Lewis, and T. G. Murphy. Simultaneous quadratic forms. *Amer. J. Math.*, 84:110–115, 1962.

[4] R. Brauer. A note on systems of homogeneous algebraic equations. *Bull. Amer. Math. Soc.*, 51:749–755, 1945.

[5] A. Cafure and G. Matera. Improved explicit estimates on the number of solutions of equations over a finite field. *Finite Fields Appl.*, 12(2):155–185, 2006.

[6] V. B. Dem'yanov. On cubic forms in discretely normed fields. *Doklady Akad. Nauk SSSR (N.S.)*, 74:889–891, 1950.

[7] V. B. Dem'yanov. Pairs of quadratic forms over a complete field with discrete norm with a finite field of residue classes. *Izv. Akad. Nauk SSSR. Ser. Mat.*, 20:307–324, 1956.

[8] H. Hasse. Darstellbarkeit von Zahlen durch quadratische Formen in einem beliebigen algebraischen Zahlkorper. *J. Reine Angew. Math.*, 153:113–130, 1924.

[9] D. B. Leep and C. C. Yeomans. The number of points on a singular curve over a finite field. *Arch. Math. (Basel)*, 63(5):420–426, 1994.

[10] D. J. Lewis. Cubic homogeneous polynomials over p-adic number fields. *Ann. of Math. (2)*, 56:473–478, 1952.

[11] D. J. Lewis and S. E. Schuur. Varieties of small degree over finite fields. *J. Reine Angew. Math.*, 262/263:293–306, 1973. Collection of articles dedicated to Helmut Hasse on his seventy-fifth birthday.

[12] W. M. Schmidt. *Equations over finite fields. An elementary approach.* Lecture Notes in Mathematics, Vol. 536. Springer-Verlag, Berlin, 1976.

[13] W. M. Schmidt. Simultaneous p-adic zeros of quadratic forms. *Monatsh. Math.*, 90(1):45–65, 1980.

[14] T. A. Springer. Some properties of cubic forms over fields with a discrete valuation. *Nederl. Akad. Wetensch. Proc. Ser. A.* **58** = *Indag. Math.*, 17:512–516, 1955.

[15] G. Terjanian. Un contre-exemple à une conjecture d'Artin. *C. R. Acad. Sci. Paris Sér. A-B*, 262:A612, 1966.

[16] E. Warning. Bemerkung zur vorstehenden Arbeit von Herrn Chevalley. *Abh. Math. Sem. Hamburg*, 11:76–83, 1935.

[17] T. D. Wooley. Artin's conjecture for septic and unidecic forms. *Acta Arith.*, 133(1):25–35, 2008.